

Food TDM use case

Title: Efficient sharing and discovery of foodborne diseases information

Commercial use (Y/N): N

Related project (if any – please specify funder and project name): Global Food Safety Partnership

Author: Giannis Stoitsis, Nikos Marianos and Nikos Manouselis

Organization: Agro-Know

Date: 22-12-2014

Use case description – Describe the **problem** you are trying to address, the **research area** it affects, and the **TDM** application used. Please state all participants, whether and how the results were subsequently used.

The problem

According to WHO foodborne and waterborne diseases kill an estimated 2.2 million people annually, most of whom are children. Diarrhoea is the acute, most common symptom of foodborne illness, but other serious consequences include kidney and liver failure, brain and neural disorders, reactive arthritis, cancer and death. The nonsufficient discovery of foodborne diseases and alerts costs has important implications for health and economy. In Europe 1.2 million cases of foodborne diseases are reported annually, leading to 350,000 hospitalisations and 5000 deaths. The estimated economic cost to Europe is 117 billion Euros per year. Most other countries in other continents similarly have economic losses. The foodborne diseases problem is not focused at the national level but also at the international level, as outbreaks involving multiple countries are becoming more common with the ever-increasing global movement of food and people. The early detection of outbreaks and the extraction of conclusions from the analysis of the alerts has become a major challenge in the area.

Currently, decision-makers in both the public and private sectors, food scientists, microbiologists and epidemiologists that work on food safety topics cannot take full advantage of all the existing data for foodborne diseases mainly for two reasons a) part of the information remains unstructured and still closed in internal databases and b) the information is stored in custom and non standard schemas and thus is not shared globally in an interoperable way.

The problem of discovery of information for foodborne diseases, food alerts, outbreaks and recalls has been stressed as a problem with major importance by global initiatives like the Global Food Safety Partnership¹ (GFSP) that works on improving the capacity building for food safety. Efficient access to and discovery of information regarding foodborne diseases will facilitate

- a) the analysis and interpretation of surveillance data by researchers and the subsequent dissemination of information to all the major stakeholders in food safety
- b) the development of early warning systems for emerging foodborne hazards
- c) the generation of customized data reports that will help decision-makers to generate foodborne alerts and recalls
- d) knowledge sharing and collaboration across sectors (public, industry, research) for foodborne diseases data surveillance

¹ <http://www.worldbank.org/en/topic/agriculture/brief/global-food-safety-partnership>

- e) standardizing methodological approaches and gaining a better understanding of the global impact of foodborne disease (estimating social, economic and health impacts)

The solution

The efficient discovery of information about foodborne diseases, food alerts, recalls and outbreaks available at diverse sources can be enabled by the use of

- a) a common semantic vocabulary (e.g. FoodEx2²) that will operate as the backbone for harmonizing the information. As the amount of information is not possible to be manually annotated with such multilingual semantic vocabulary, efficient text mining methods are needed that will automatically extract vocabulary/ontology terms from such information
- b) text mining tools that will extract structured data from trusted web sources containing information about foodborne diseases, outbreaks and recalls
- c) text mining tools that can extract the secondary data from publications and reports such as tables with results, images and genetic information (DNA of bacteria)
- d) tools that will transform the annotated information for foodborne diseases to linked data, maximizing the processing of the information by machines.
- e) of text mining tools that will enable the automatic linking of food outbreaks' reports with food recall data.

Addressing the above data challenges will allow scientists to find foodborne diseases information in order to analyse and correlate the data and to create notifications about food safety issues. Apart from the data and technological enablers the most important prerequisite for applying state-of-the art text analysis techniques is the availability and openness of information regarding the foodborne diseases, outbreaks, food alerts and recalls.

Content sources – List the targeted content sources and types of content to used.

- Publications, reports, classification and monitoring data available by the **European Food Safety Authority** (<http://www.efsa.europa.eu/>).
- The **Rapid Alert System for Food and Feed** System (http://ec.europa.eu/food/safety/rasff/index_en.htm) and specifically its machine interfaces such as the SPARQL end point that can be used to perform sophisticated queries on historical data about food alerts, recalls and measures taken by the national authorities.
- The **Center for Science in the Public Interest** (<http://www.cspinet.org/>) and specifically **a)** the outbreaks database (<http://www.cspinet.org/foodsafety/outbreak/pathogen.php>) that tracks foodborne illness outbreaks in US since 1997 and **b)** the database of outbreak reports (http://cspinet.org/foodsafety/outbreak_report.html).
- **The Centers for Disease Control and Prevention** (<http://www.cdc.gov/>) and specifically the database of the foodborne outbreak reports.
- Information about food safety published by the **International Food Policy Research Institute** (<http://www.ifpri.org>) and specifically machine interfaces such as the SPARQL end point that the institute is providing (<http://data.ifpri.org/sparql>).
- Reports published by **Food Safety International** (<http://www.safefoodinternational.org/>).
- Social Media such as **Twitter** that can be used to identify foodborne disease incidents

² <http://www.efsa.europa.eu/en/supporting/pub/215e.htm>

reported directly by food consumers.³

- Data about search queries related to food safety incidents from search engines like **Google** and **Yahoo**.³

Targeted users – Describe end users, their number and expertise.

- **Decision-makers** working in international organizations like the European Food Safety Agency, WHO, Food and Agricultural Organization of United Nations (FAO) and National Authorities that face the challenge to define food safety policies, recommendations and measures for foodborne data surveillance
- **Epidemiologists** working on the processing of information for foodborne diseases, outbreaks and alerts trying to identify correlations with food recalls.
- **Microbiologists** working on identifying microbiological hazards which are caused by pathogenic (disease causing) microorganisms. They need to have direct access to data from the microbiological tests.

Impact – Describe all possible impacts of the use case. Specify any cross-border, societal and economic effects (if possible state any monetary benefits and market advantages).

Cross-border

Cultural and demographic factors, as well as increased mobility, have resulted in major epidemiologic shifts in foodborne disease during recent decades. Previous outbreaks of foodborne disease were smaller and limited in scope, and more often originated in the home. Nowadays, many more people dine outside the home and travel more extensively. As a result, more than 80% of foodborne disease cases occur from exposures outside the home. International collaborations, information technologies and data/knowledge sharing are required to address this challenge. Efficient discovery of foodborne diseases can facilitate the cross-border collaboration and exchange of information among researchers and policy-makers.

Societal

Foodborne diseases impose a substantial burden on health-care systems, trade and tourism, markedly reduce economic productivity and threaten livelihood. In Europe 1.2 million cases of food borne diseases are reported annually, that lead to 350,000 hospitalisations and 5000 deaths. According to WHO Diarrhoea is the most common foodborne illness caused by pathogens, but other serious consequences include kidney and liver failure, brain and neural disorders, reactive arthritis and death. Chemical food contamination may cause non-communicable diseases, in particular cancer, and can also affect reproductive health and the immune system. Furthermore, the impact of food safety in European consumers' confidence is also very important. Recent studies have concluded that consumer confidence in European food safety has declined in recent years. Improved sharing and discovery of the foodborne diseases information can contribute in the production of safer foods and subsequently in increasing consumer confidence.

Economic

As already mentioned the estimated economical cost in Europe from foodborne diseases is 117 billion Euros per year. Much of these losses represent lost markets, loss of consumer demand,

³ <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3785982/>

litigation and company closures. Improving the dissemination of information regarding food safety issues of raw material could help companies in the agrifood sector to recall food products and to limit or even prevent economic losses.

Constraints – *Describe any legal, technical, economic, societal, organizational, cultural, multi-lingual or other limitations and you have encountered.*

The main constraints that limit the discovery of foodborne diseases information are the following:

- *Legal:* No clear licensing framework exists for sharing the information of foodborne diseases. The legal framework for the specific type of information should be developed respecting the sensitivity of the foodborne diseases information when needed. Currently, important international organizations like the World Health Organization (WHO) do not have adopted open access policy for all the information products related to food safety.
- *Data availability:* Data from food supply chain about the chemical substances and pesticides used during cultivation and production of the food are not currently available as open data
- *Multilinguality.* The food alerts are disseminated in all EU languages. The text mining algorithms that will be used to automatically annotate the foodborne information should support at least European languages and thus big amount of language resources are needed for the training of text mining algorithms. Also the metadata of the foodborne disease data needs to be translated into multiple languages.

Comments – *Other comments and any recommendations you may have.*